UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO INFORMÁTICA

Sistema Automático para Recolha de OSINT e Integração com Plataforma de Threat Intel

Diogo António Cardoso Dionísio

Mestrado em Engenharia Informática
Especialização em Arquitetura, Sistemas e Redes de Computadores

Versão Pública

Trabalho de Projeto orientado por:
Professor Doutor Luís Manuel Ferreira Fernandes Moniz

2019

Agradecimentos

Quero dedicar este espaço a quem tornou possível, para além da minha pessoa, a conclusão desta dissertação.

Antes de mais quero agradecer ao meu Pai, Mãe, Vítor, Ana e aos meus irmãos Simão, Rita e Tatiana por me terem facultado tudo o que era necessário e toda a confiança neste longo processo académico, sem eles não teria a oportunidade de ter chegado onde cheguei. Um especial agradecimento às minhas irmãs por não me deixarem dormir ao fim-de-semana e ao meu irmão por me ter tirado a consola, foram sem dúvida peças fundamentais para chegar ao final deste percurso.

À minha namorada, confidente, ouvinte e conselheira Ana Catarina Travessa, por me ter acompanhado e ajudado ao longo de toda esta caminhada, principalmente durante o desenvolvimento deste projeto.

À família Ribatejana por todos os jogos noturnos, cafés, passeios e acima de tudo por toda a amizade que existe. Em especial quero agradecer ao Miguel Andrade por ser a pessoa mais desagradável que existe, no entanto é e será sempre o meu principal conselheiro e parceiro de desenvolvimento de *software* sem fins lucrativos.

Aos meus colegas e amigos do Departamento de Informática que sem dúvida contribuíram para todo este percurso académico. Alex, Miguel, Patination, Robin dos Bosques, Tio Jorge e ao Because Prontos. Um especial agradecimento ao Diogo Gonçalves e Gonçalo Domingos por serem os parceiros de guerra nos trabalhos de grupo do curso de Engenharia Informática, seremos sempre o grupo ”Masters of Everything”com o melhor jogo *mobile* - ”EarthWorld”! Sem dúvida que todos eles serão amizade para a vida.

Um grande especial agradecimento ao Nuno Matos por depositar em mim toda a confiança, colaboração e transmitir todo o seu conhecimento para que pudesse crescer como pessoa e como profissional. Ao Professor Doutor Luís Manuel Ferreira Fernandes Moniz por toda paciência, tolerância e ajuda no desenvolvimento desta dissertação ao longo do ano. À fantástica instituição LAYER8 por me ter acolhido e a todos os seus elementos que, de alguma forma, contribuíram para o desenvolvimento deste projeto.

Por fim quero agradecer à Faculdade de Ciências da Universidade de Lisboa por ter tido a oportunidade de adquirir conhecimentos em Engenharia Informática e me transmitir valores que levo comigo para a vida.

Aos meus pais e irmãos
Ao meu avô e avó

Resumo

As novas ameaças surgem permanentemente, atormentando a nova era tecnológica. A cada uma delas, estão inerentes caraterísticas específicas, bastante variadas entre si, tal como formas distintas de comportamento. Cada comportamento e característica de uma ameaça possuí indicadores associados a ele e são chamados de *Indicators of Compromise* (IOCs) que podem ser IPs, domínios, *hashes*, entre outras. Os IOCs são considerados artefactos forenses e são utilizados como sinal de que um sistema foi comprometido ou infetado por um determinado *software* malicioso. Assim, é necessário que haja exatidão na recolha de informação para encontrar diferentes IOCs. Deste modo, é importante a existência de standards e de plataformas autónomas para a partilha e recolha de informação de *cyber* segurança de forma a ajudar as organizações a tornarem-se cada vez mais resilientes a novas ameaças.

As plataformas autónomas atualmente existentes para a recolha e análise de indicadores são sistemas ainda muito independentes entre si, o que dificulta a automatização do processo na integração com as plataformas de *Security Information and Event Management* (SIEM) para a correlação de eventos. A existência de um sistema de *Threat Intel* capaz de recolher vários indicadores de diferentes fontes abertas (*Open Source Intelligence* ou OSINT) e de análises manuais internas dos analistas, é uma mais valia que facilita esta integração com diferentes infraestruturas/redes.

Este projeto teve como base fundamental a plataforma *Malware Intelligence Sharing Platform* (MISP). O MISP funcionará como uma base de dados de indicadores, de forma a que a mesma seja consultada para a procura de IOCs, facilitando não só a prevenção, como também a fase de análise de um *malware*. Para além disso, na recolha de informação será usada informação real recolhida por inúmeras vias, entre elas: SIEM; *Feeds OpenSource*; e ferramentas internas à LAYER8, não esquecendo também as análises manuais que são de facto uma mais valia. Por último, na fase da análise de IOCs e com vista a automatizar este processo, será realizada uma combinação de vários sistemas de análise, na qual os mesmos serão ligados à plataforma MISP.

Em suma, este projeto será assim um laboratório de análise de ameaças, recolha e partilha de IOCs, dando a estes um nível de confiança adequado para serem usados posteriormente.

Palavras-chave: *Threat Intel*; MISP; Cortex; SIEM; SOC; IOC

Abstract

There are constantly new threats surging, tormenting the new technological era. To each of these, there are inherent specific characteristic, which can widely vary from one another, as well as their distinctive behaviours. Each behaviour and characteristic of a threat carries associated indicators denominated Indicators of Compromise (IOCs) that can be IPs, domains, hashes, in-between others. The IOCs are considered forensic artefacts and used as a symbol that a system has been compromised or infected by a certain malicious software. Therefore, it is necessary that the data recovery is accurate so that different IOCs may be identified.

In this way, the existence of standards and autonomous platforms to share and collect cyber security information in order to help organisations becoming more resilient to new threats, becomes crucial.

The currently existent autonomous platforms for data collection and analysis of indicators are still very independent in between one another's, which makes difficult the automatization of the process in the incorporation with the Security Information and Event Management (SIEM) platforms, to enable event correlation. The existence of a system of Threat Intel capable of collecting multiple indicators from different open sources (Open Source Intelligence ou OSINT) and of analysing internal manuals from analysts, is an added value that eases this integration with different infrastructures/networks.

This project was fundamentally based on the Malware Intelligence Sharing Platform (MISP). The MSIP will work as a data base of indicators, in order to facilitate the same to be consulted for the search for IOCs. This will ease not only the prevention, as well as the analysis phase of a malware. Furthermore, the collection of information will use real data collected through numerous pathways, such as: SIEM; Feeds OpenSource; and LAYER8 internal tools, keeping in mind the manual analyses that are in themselves a plus. Lastly, during the analysis phase of the IOCs, and looking to automate this process, a combination of various analysis systems will be conducted, in which the mentioned will be connected to the MISP platform.

Ultimately, this project will be a threat analysis laboratory, collecting and sharing IOCs, providing those with an adequate trust level to enable their future use.

Keywords: *Threat Intel*; MISP; Cortex; SIEM; SOC; IOC

Conteúdo

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

Este capítulo introduz o projeto desta dissertação onde a motivação, objetivos e contribuições serão aqui discutidas assim como também será feita a apresentação da restante estrutura deste projeto.

A maioria das ameaças à segurança das organizações provêm de códigos maliciosos ou através de engenharia social que podem ser distribuídas das mais variadas maneiras, tais como na exploração de vulnerabilidades, envio de *e-mails* maliciosos para disseminar campanhas de *phishing*, entre outros. O objetivo destas ameaças pode ser desde a infiltração de forma ilícita num sistema prejudicando deliberadamente o seu normal funcionamento ou até no roubo de informações (desde documentos a dados pessoais) de um respetivo alvo ou alvos em massa.

Esta deteção de atividades maliciosas pode ser feita através de *Indicators of Compromise* (IOCs) em que cada indicador está associado ao comportamento ou característica de uma ameaça específica. Estes podem ser, por exemplo IP de origem dos pedidos maliciosos, a *hash* de um ficheiro de *malware*, um endereço de *e-mail* de origem, entre outros. Como um IOC é um identificador de comportamentos maliciosos e devido à volatilidade dos sistemas que os atacantes utilizam, um IOC pode ter um período curto de vida.

A fim de evitar IOCs sem atividade maliciosa ao dia da consulta, é importante, que sejam analisados periodicamente e armazenados num local centralizado, dando assim contexto e confiança ao indicador consultado. Esta confiança, é crucial para que seja possível a sua utilização por analistas no SOC, de forma automática nas plataformas de *Security Information and Event Management* (SIEM) e equipamentos de segurança na tomada de ações de prevenção quando este é encontrado nas infraestruturas/redes das organizações.

É por isso importante que os sistemas de deteção e SIEM, possuam uma boa base de dados de IOCs associados a assinaturas de diferentes comportamentos maliciosos, de modo a que sejam detetados em tempo útil para prevenir a segurança da infraestrutura de rede de uma organização. Assim torna-se evidente a importância da recolha, partilha e análise de indicadores num local centralizado na segurança das redes em tempo útil.

1.1 Motivação

Atualmente as listas de IOCs disponíveis *online* pertencem a diferentes plataformas como é o caso do AlienVault[27] e Tallos Intelligence[11]. Estas listas de IOCs reputados são muito à base de partilha de IPs ou domínios que foram referenciados por outros sem contexto sobre o mesmo. Por essa razão existe o problema de não se saber a razão de um certo indicador estar reportado nessa lista, uma vez que não existe qualquer contexto nem local onde foi inicialmente visto com comportamento malicioso, o que pode levar a certas lacunas na confiança do IOC para a sua utilização.

Outra razão que abala a confiança dos IOCs pertencentes a estas listas é a volatilidade dos sistemas que os atacantes utilizam. Essa volatilidade pode fazer, por exemplo que um indicador que foi referenciado no passado devido a comportamento malicioso, ao dia de hoje esse indicador já não pertencer ao atacante e por isso ser um artefacto legítimo. Isto poderá desenvolver falsas conclusões de um analista levando a bloqueios mal realizados a um certo IP ou domínio que são realmente legítimos. Por esta razão é bastante importante que os IOCs recolhidos destas listas tenham confiança suficiente para a sua utilização na segurança proativa das infraestruturas.

É por isso fulcral que os IOCs guardados na base de dados possuam algum contexto, ou algum tipo de métrica baseado numa análise, de forma a garantir a confiança na utilização do indicador para correlação de eventos na plataforma de SIEM assim como nos dispositivos de segurança da infraestrutura.

1.2 Objetivos

O objetivo deste projeto de dissertação é a integração de um sistema capaz de recolher, armazenar e analisar IOCs de forma a aumentar a confiança nos indicadores armazenados, permitindo desta forma que o sistema tenha capacidade de reconhecer e atuar contra ameaças em tempo útil através dos IOCs.

Assim, a recolha de IOCs é realizada através da integração da plataforma interna DNS8 ao SOC da LAYER8, listas de comunidades *online*, SIEM ou manualmente por analistas para um repositório central.

Os indicadores recolhidos e armazenados no repositório, são atualizados com análises automáticas regulares. Estas análises são realizadas com o objetivo de evitar o uso dos indicadores que já não possuem comportamento malicioso associado. Desta forma, a sua utilização é possível de forma automática no SIEM e nos dispositivos de segurança na infraestrutura de rede com o objetivo na tomada de ações defensivas em tempo útil.

1.3 Contribuições do Trabalho

O desenvolvimento deste trabalho contribuiu para:

- Aumento da segurança nas infraestruturas de rede dos clientes do SOC da LAYER8 através da sua monitorização proativa;
- Integração de dois serviços da LAYER8, SOC8 e DNS8, contribuindo para o crescimento dos serviços através da cooperação entre ambos;
- Funcionalidades que facilitam na tomada de decisões dos analistas do SOC da LAYER8;

1.4 Enquadramento Institucional

Este trabalho foi realizado no âmbito da disciplina de Dissertação/Projeto de Engenharia Informática (DPEI) da Faculdade de Ciências da Universidade de Lisboa (FCUL). A realização deste trabalho é o requisito para aprovação a esta disciplina que por sua vez é o requisito final para a conclusão do Mestrado em Engenharia Informática (MEI) na especialização de Arquitetura, Redes e Sistemas de Computadores. O trabalho foi desenvolvido numa instituição privada e externa à Universidade de Lisboa (UL), LAYER8 - Shield Domain, SA.

1.5 Estrutura do Documento

Este documento está organizado da seguinte forma:

- Capítulo 2 – Contexto do projeto de dissertação;
- Capítulo 3 – Ferramentas a utilizar neste projeto;
- Capítulo 4 – Descrição do módulo de recolha automática de OSINT;
- Capítulo 5 – Descrição do módulo de análise de indicadores;
- Capítulo 6 – Descrição da plataforma de *threat intel* desenvolvida para este projeto;
- Capítulo 7 - Integração dos módulos desenvolvidos;
- Capítulo 8 – Exemplos de casos práticos do sistema desenvolvido em ambiente real;
- Capítulo 9 - Conclusão do projeto de dissertação;

Capítulo 2

Contexto

Neste capítulo serão apresentados conceitos teóricos relevantes para o tema deste trabalho de dissertação. Esta apresentação tem como objetivo a introdução e o entendimento dos temas aqui discutidos. Como ponto de partida, será apresentado o conceito teórico sobre temas como a segurança da informação, OSINT, indicadores de compromisso, *Thread Intel* e também como estes estão relacionados com um centro de operações de segurança (SOC).

2.1 Segurança da Informação

O uso da informação está cada vez mais associado diretamente aos processos de negócio das organizações. Essa informação pode ser muito variada, podendo ir de informação confidencial de utilizadores até chamadas a serviços aplicacionais entre máquinas. Qualquer tipo de interação com estes processos que perturbe o seu normal comportamento pode resultar em grandes perdas indesejáveis para uma organização, o que justifica o aumento no investimento em segurança informática. Neste contexto, para além do investimento em tecnologia, a organização necessita de capacitar os seus recursos humanos através de formação sobre como atuar, de forma a evitar comportamentos de risco e tendências que possam fragilizar a sua segurança pessoal e da organização.

A segurança da informação está relacionada com a proteção de um conjunto de dados sendo o objetivo principal, segundo a tríade CIA (*Confidentiality*, *Integrity* e *Availability*), garantir não só a confiabilidade e integridade, mas também a disponibilidade da informação. Além de proteger a informação, garante que todos os incidentes de segurança são tratados segundos os procedimentos alinhados com as políticas internas em vigor e caso voltem a acontecer, serão detetados e resolvidos de forma mais célere, reduzindo o potencial impacto nas infraestruturas.

2.2 Indicadores de Compromisso (IOC)

Aquando a existência de um crime, tende-se em procurar por pistas de forma a ser possível responder a uma série de perguntas importantes, tais como: "Porquê?"; e "Como aconteceu?". Estas questões, não só irão ajudar na compreensão dos acontecimentos, como também irão permitir correlacionar o mesmo com os outrora existentes, ajudando na sua resolução.

A ideia dos IOCs é exatamente a mesma de como se um crime se tratasse, sendo neste caso um IOC uma pista de um crime. Ou seja, esta recolha de IOCs segue a mesma metodologia das recolhas de provas de um crime. Mesmo que um atacante tente minimizar o seu rasto, existe sempre uma pegada associada ao seu ataque.

Um IOC segundo [15] é considerado um dado forense que servirá para identificar se um sistema foi comprometido ou infetado por um atacante com por exemplo um *software* em específico. Assim sendo, um IOC pode ser endereços de *e-mail*, IPs, URLs, *Hashes*, chaves de registo e tudo o que seja possível retirar de um certo evento de origem maliciosa. Dando exemplo de um caso prático onde ocorreu tentativas abruptas de *login* na conta de um utilizador de uma organização, o IOC recolhido nesta situação poderá ser por exemplo o IP em que os pedidos tiveram origem nestas tentativas de *login*.

A partilha de indicadores de um certo acontecimento com outras organizações é extremamente importante de modo a que estas possam prevenir a defesa dos seus sistemas perante um ataque igual à sua organização. Tendo o conhecimento dos indicadores de que um certo ataque existe, poderão agir proativamente nas suas medidas defensivas contra esses ataques, como por exemplo, na deteção da comunicação de uma máquina na rede da organização com um *malware* que utiliza um certo IP para a sua comunicação. Essa deteção é feita através dos logs dos sistemas que podem ajudar a identificar a atividade potencialmente maliciosa numa infraestrutura através de IOCs.

2.3 *Open Source Intelligence* (OSINT)

A sigla OSINT provém de *Open-Source Intelligence* é o termo dado à recolha de informação a partir de dados públicos e recolhidos de várias fontes de informação [28]. Essa informação está disponível em fontes públicas, incluindo a informação pública da *internet* e por isso é possível coletar informação (ou inteligência) seja ela sobre pessoas, organizações, *tweets* ou até de uma fotografia [28].

No âmbito da segurança informática, este tema possui uma grande importância pela capacidade de ajudar na predição, prevenção e deteção de possíveis incidentes e ameaças de segurança dirigidas a uma organização. As OSINT tornam-se muito valiosas quando correlacionadas com os eventos gerados pelos equipamentos de rede e/ou máquinas pessoais ativos numa organização, assim como os eventos gerados a nível aplicacional de modo a agir proativamente perante novos incidentes. Este trabalho de correlacionar é

desempenhado de forma automática essencialmente via SIEM e de forma manual por analistas de SOC.

2.4 *Threat Intelligence*

A *Threat Intelligence* é a previsão do conhecimento baseado em evidências sobre ameaças existentes ou potenciais com destino à sua organização. Este conhecimento é adquirido essencialmente com a informação recolhida de *feeds* OSINT como também de qualquer outro tipo de eventos, esta informação é usada para prevenir e identificar ciber ataques que procuraram tirar vantagem de recursos valiosos para uma organização. Assim a *threat intelligence* permite adquirir vantagem sobre eventuais ataques às organizações, ajudando a construir mecanismos efetivos de defesa para que o risco ou a infeção seja minimizado e mitigando da forma mais eficaz possível.

As soluções de inteligência de ameaças coletam muita informação sobre todo o tipo de ameaças e atores emergentes bem como informação existente de diversas fontes incluindo as OSINT. Após esta recolha de inúmeras fontes, esta informação é analisada e filtrada de modo a produzir *feeds* de *threat intel*. Com estas *feeds*, os sistemas autónomos podem utilizá-las com a finalidade de proteger proativamente uma organização bem como alertar as equipas responsáveis.

Assim sendo, o principal objetivo da *threat intelligence* é manter as organizações informadas dos riscos de ameaças avançadas persistentes, *zero-day threats* e *exploits* de forma a proteger as organizações contra os mesmos.

2.5 *Security Operations Center* (SOC)

Uma sala espaçosa com equipamentos de topo, ecrãs gigantes e muitos Engenheiros especializados, não é necessariamente um requisito para formar um Centro de Operações de Segurança. Um SOC, de acordo com [32] é composto por uma equipa de analistas de segurança organizados para detetar, analisar, responder, reportar e prevenir incidentes de segurança cibernética.

Um SOC pode fornecer uma variedade de serviços a um conjunto de clientes e cada cliente é referido como *constituency*, que de acordo com [32] pode ser definido como um conjunto restrito de utilizadores, sites, ativos, redes e organizações. Os serviços fornecidos por um SOC podem variar desde a sensibilização dos utilizadores para alguns dos riscos que estes estão expostos no seu dia-a-dia (*security awarness*) a avaliações de vulnerabilidades como a identificação, quantificação e priorização de vulnerabilidades existentes.

Contudo, um SOC com uma boa arquitetura é capaz de fornecer uma visibilidade completa e precisa da infraestrutura monitorizada, o que resulta numa postura de segurança

mais forte. Do ponto de vista dos analistas, ter um local central com informação significativa da monitorização da infraestrutura permite uma deteção de ameaças mais rápida e eficiente.

De modo a que uma equipa de analistas consiga aproveitar ao máximo este local central de informação a fim de reduzir os riscos de segurança numa organização, um SOC necessita de ter pessoas com formação especializada e certificada, processos e tecnologia. Estes analistas do SOC devem ser altamente versáteis a trabalhar com os mais diversos cenários, uma vez que são descobertos diariamente diferentes tipos de vulnerabilidades e novas ameaças. Isto obriga a que a equipa de analistas trabalhe sobre pressão e sejam extremamente versáteis com qualquer tipo de ameaça.

Um sistema de monitoramento de segurança eficiente incorpora de forma continua dados de diferentes fontes, como por exemplo dados de sistemas de segurança colocados na infraestrutura ou até mesmo de computadores pertencentes à infraestrutura da organização. Como esta agregação de dados acontece antes e durante um incidente de segurança, a equipa pode começar a usar de imediato esse sistema de monitorização como uma ferramenta de deteção sendo também possível usá-la como um ferramenta de investigação uma vez que é possível verificar todos os passos que aconteceram antes de se dar o incidente.

2.5.1 Logs

Os logs são indicadores cruciais para se identificar o que está a acontecer numa infraestrutura de rede de uma organização. São os logs que ajudam os analistas a correlacionar por exemplo comportamentos de rede, fornecendo informações muito valiosas sobre diferentes tipos de problemas de segurança numa infraestrutura. Estes logs para além de ajudarem a evitar que uma organização possua uma visão limitada sobre o que está a acontecer na sua infraestrutura, permite também aos especialistas de segurança a verificarem comportamentos maliciosos com intuito de prejudicar a mesma.

Um log de acordo com [17] é a forma mais básica de informação que um sistema pode gerar. Estes logs são geralmente produzidos como uma forma de auditoria, dando indicações de que algo aconteceu erradamente podendo até, correlacionando todos os logs, fornecer informações sobre como surgiu o erro. Podem ser gerados por um serviço aplicacional, sistema operativo, equipamento de rede etc. e regista toda a informação básica sobre algo que acabou de acontecer. Esta informação por mais básica que seja, pode ser informação muito valiosa que ajuda no diagnóstico de problemas, desde serviços que não executaram corretamente a comunicações maliciosas entre um *host* da rede e um serviço na web, a figura 2.1 representa um log de uma *firewall*.

Atualmente, os modernos sistemas de gestão de incidentes e eventos de segurança (SIEMs) funcionam correlacionando estes logs de diferentes fontes de modo a evitar ou a identificar incidentes de segurança. Um SIEM usufruindo desta basta coleção de logs,

pode correlacioná-los para por exemplo alertar uma equipa do SOC avisando de que um utilizador específico está a realizar atividade incomum na sua organização.

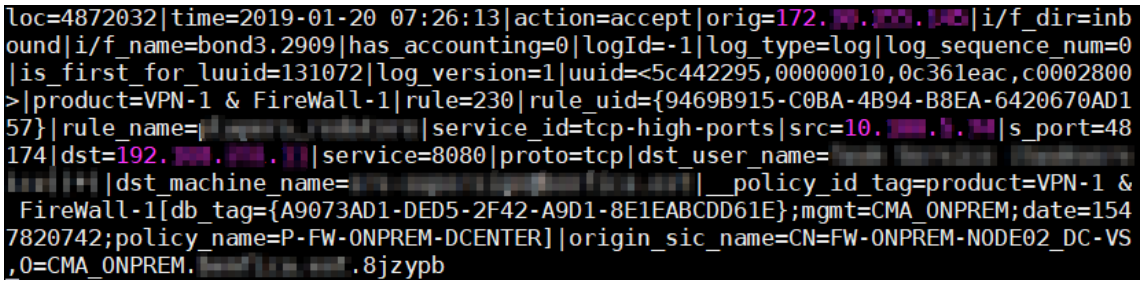

Figura 2.1: Representação de um log de uma *firewall* Checkpoint

2.5.2 SIEM - *Security Information and Event Management*

Uma plataforma de SIEM tem origem em dois conceitos completamente diferentes mas cada um criado com intuito de resolver as suas necessidades. Esses dois conceitos que deram origem ao SIEM são: *Security Information Management* ou Gestão de Informação de Segurança (SIM) e *Security Event Management* ou Gestão de Eventos de Segurança (SEM). O SIM foi feito com a necessidade de resolver a capacidade de análise em tempo real, com vista a melhorar a resposta a incidentes enquanto o SEM foi feito orientado ao armazenamento de longa duração e análise histórica dos eventos de modo a suportar atividades de análise forense.

O termo de *Security Information Event Management* ou Gestão de Informação de Eventos de Incidentes de Segurança (SIEM), que foi batizado por Mark Nicolett e Amrit Williams do Gartner em 2005 [14], descreve um SIEM como uma plataforma com a capacidade de coletar, analisar e apresentar informações de uma infraestrutura ou de equipamentos de rede. Portanto um SIEM é uma solução híbrida que combina as ferramentas SIM e SEM numa única plataforma com o objetivo, segundo [18], de ajudar as organizações ou às suas equipas de analistas a responder a incidentes de segurança de uma forma rápida bem como na organização da grande quantidade de informações recolhidas pelo SIM.

Assim o SIEM permite o gerenciamento e correlação de eventos gerados por diversas aplicações de segurança tais como *firewalls, proxies*, *Intrusion Prevention System* (IPS), *Web Application Firewall* (WAF) entre outros. Estes eventos são então coletados, normalizados, armazenados e correlacionados pelo SIEM. Tendo assim inúmeros eventos armazenados possibilita uma rápida identificação e resposta aos incidentes, nalguns casos esta resposta até pode ser automática por parte do sistema. Mantendo um histórico de todas as interações com a infraestrutura, sendo também bastante útil para investigações forenses.

2.6 Conclusão

O conteúdo teórico apresentado torna-se relevante para o desenvolvimento do trabalho uma vez que suporta os pressupostos enquadrados com os objetivos do projeto. A *threat intelligence* através da correlação de eventos nos SIEMs com os indicadores de compromisso recolhidos nas OSINT faz com que seja possível reconhecer ameaças nos sistemas das organizações, permitindo ao SOC atuar sobre estas ameaças de modo a aumentar a segurança da informação das organizações. Assim, é nesta base que se iniciou o projeto que se encontra desenvolvido nos capítulos seguintes.

Capítulo 3

Trabalho Relacionado

Neste capítulo será apresentado o trabalho atualmente existente e que é fundamental para o tema deste trabalho. O objetivo deste capítulo é dar a conhecer o trabalho que atualmente tem vindo a ser desenvolvido relacionado com este tema.

3.1 MISP (*Malware Intelligent Sharing Platform*)

O MISP é uma plataforma gratuita e *open source*, desenvolvida pelo *Computer Incident Response Center Luxembourg* (CIRCL)[4] com o objetivo de responder às exigências desta equipa e também poder ser uma ferramenta de trabalho no dia-a-dia desta equipa. O CIRCL foi criado pelo governo do Luxemburgo com o objetivo de oferecer uma resposta sistemática às ameaças e incidentes de *cyber* segurança. O MISP tem como objetivo recolher, guardar, partilhar e correlacionar indicadores de compromisso de informações sobre (*thread intel*). Estas informações de *Thread Intel* incluem indicadores de segurança cibernética, ataques direcionados, ameaças, vulnerabilidades ou até mesmo informações de anti-terrorismo.

O MISP não é a única plataforma desenhada para efeito de partilha de indicadores. No entanto, este projeto está direcionado para esta plataforma, pois como referido acima, esta é uma ferramenta *open source*, desenhada por analistas de incidentes e de engenharia reversa de malware, profissionais de tecnologias de informação (IT) e de segurança cibernética, com o objetivo de que esta os ajude nas tarefas diárias de forma eficiente.

3.1.1 Base do MISP

A ideia principal da plataforma MISP é criar uma base de dados de IOCs. Esta servirá para agrupar indicadores que estejam relacionados entre si, isto é, num certo ataque em que foram identificados o *e-mail* “mail@ioc.malware” e o IP “154.69.68.5”, é necessário manter estes dois IOCs juntos, pois existe uma relação entre ambos na altura da realização do ataque.

No MISP um ataque é representado por um termo geral de *Event* e os IOCs identificados serão os atributos de um evento específico. Esta representação proporciona assim uma fácil de identificação e correlação entre eventos graças aos seus atributos, por exemplo, um IP ou *e-mail* estarem associados a mais que um evento. Esta representação é importante não só para relacionar atributos de eventos, como também permitir a identificação no tempo de quando é que um determinado IP esteve em funcionamento com finalidades maliciosas. Isto é muito importante porque um IP dinâmico ao dia de hoje poderá ser malicioso, mas amanhã já não o ser.

Outra questão importante no MISP é o facto de que uma vez que este é um projeto *open source* significa que qualquer pessoa poderá correr uma instância MISP isolada, no entanto é possível interligar várias instâncias MISP e sincronizar eventos entre elas. Sendo também possível garantir a confidencialidade de certos atributos ou eventos, o que é bastante importante pois assim é possível restringir certo tipos de eventos com informações confidenciais com outras instâncias e/ou utilizadores da instância.

Além disso por ser uma plataforma *open source*, muitas ferramentas adicionais têm sido criadas com o intuito de facilitar o seu desenvolvimento e crescimento. Exemplo disso é a biblioteca PyMISP[3] para a linguagem de programação Python que permite interagir facilmente com a instância para que haja automação no uso das funcionalidades do MISP.

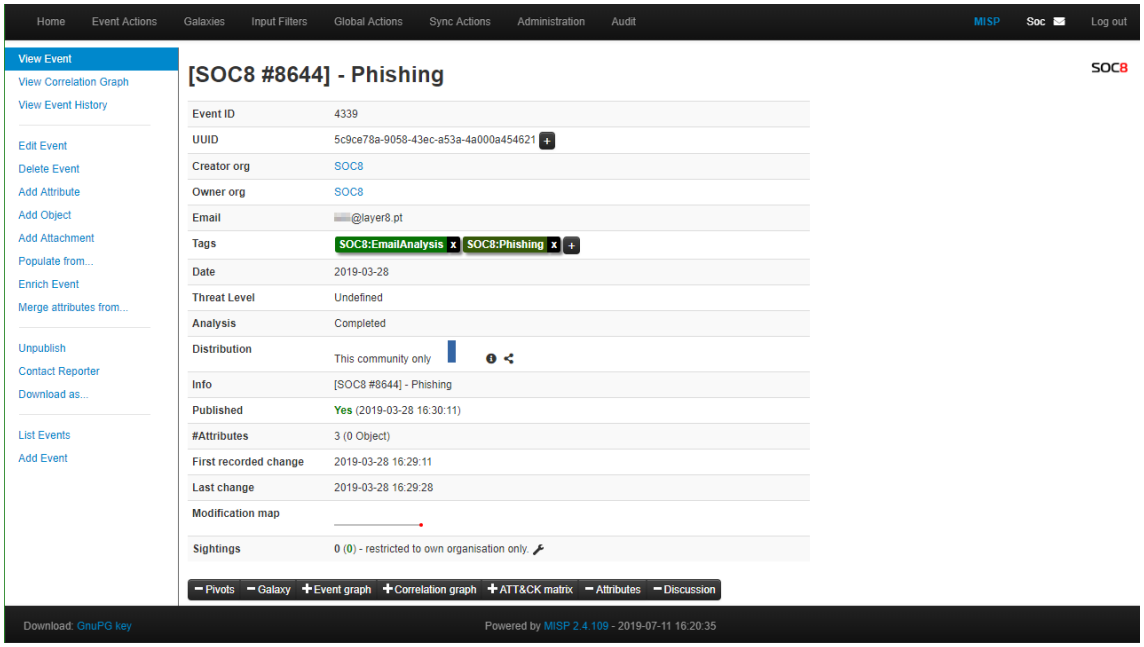

Figura 3.1: Informação do evento MISP

3.1.2 Guia de Partilha Traffic Light Protocol (TLP)

O Traffic Light Protocol(TLP) foi criado para facilitar a partilha de informação definindo níveis de autorização de divulgação. O TLP tem um conjunto de definições usadas para garantir que informações confidenciais sejam partilhadas com o público alvo apropriado e por isso possui quatro cores que indicam o nível de permissão de partilha da informação. Segundo a FIRST[8] o TLP possui as cores vermelho, amarelo, verde e branco. Quais queres outras cores não são consideradas pela FIRST. Este esquema TLP é também usado pelo Centro Nacional de Cibersegurança (CNCS) [5] e por isso adotado pela empresa LAYER8.

Cor	Permissão
RED	Apenas para os participantes, não é para divulgação
AMBER	Restrita às organizações participantes, divulgação limitada
GREEN	Restrita à comunidade, divulgação limitada
WHITE	Divulgação não limitada

Tabela 3.1: Representação das permissões do protocolo TLP

3.1.3 Estrutura ao MISP

A página principal do MISP é uma lista com todos os últimos eventos adicionados à instância como pode ser observado na figura 3.2.

Cada evento possui TAGs de modo a que seja possível uma melhor organização nessa base de dados de eventos, bem como tornar mais fácil a sua pesquisa de eventos por categorias de eventos.

Quando se entra num evento, figura 3.1, é possível encontrar toda a informação associada, nomeadamente uma lista de atributos de um evento específico 3.3.

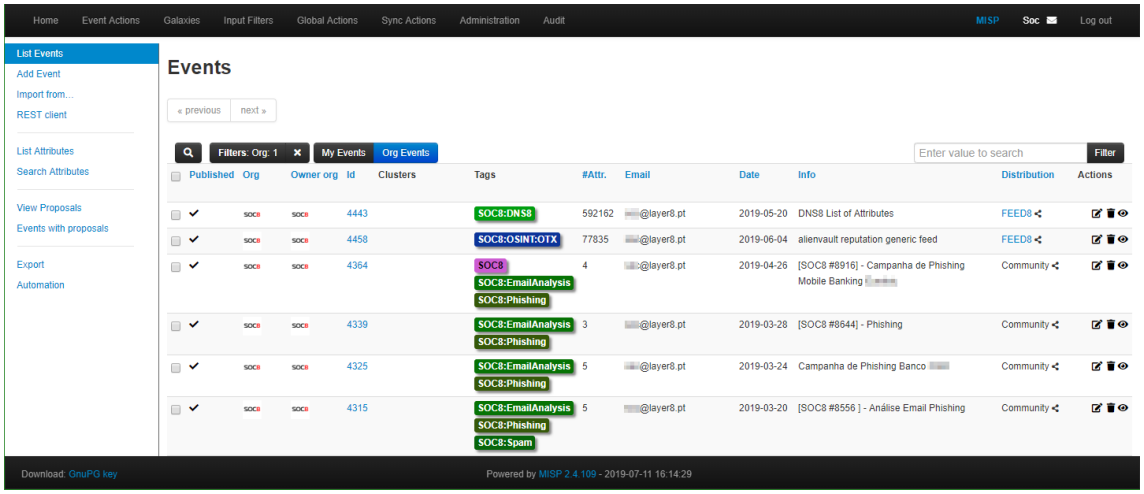

Figura 3.2: Lista de Eventos no MISP

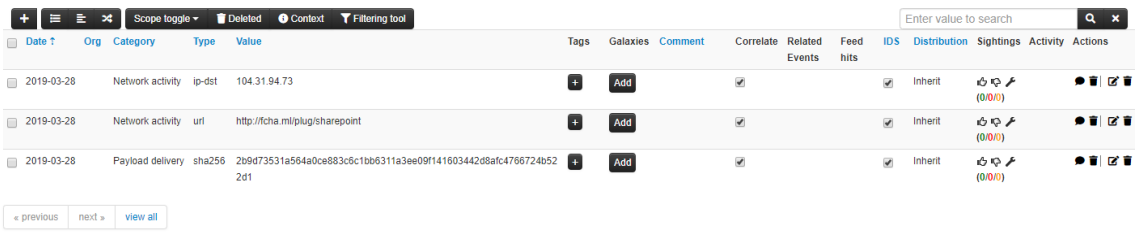

Figura 3.3: Atributos do Evento

Por último, é possível também visualizar um grafo com todas as correlações existentes nessa instância MISP. Essa correlação associa os atributos desse evento em questão com os atributos de todos os outros eventos existentes, tornando assim possível, verificar se há outras ocorrências desse mesmo atributo, facilitando assim o analista a procurar informação. A figura 3.4 mostra a correlação de atributos na plataforma MISP.

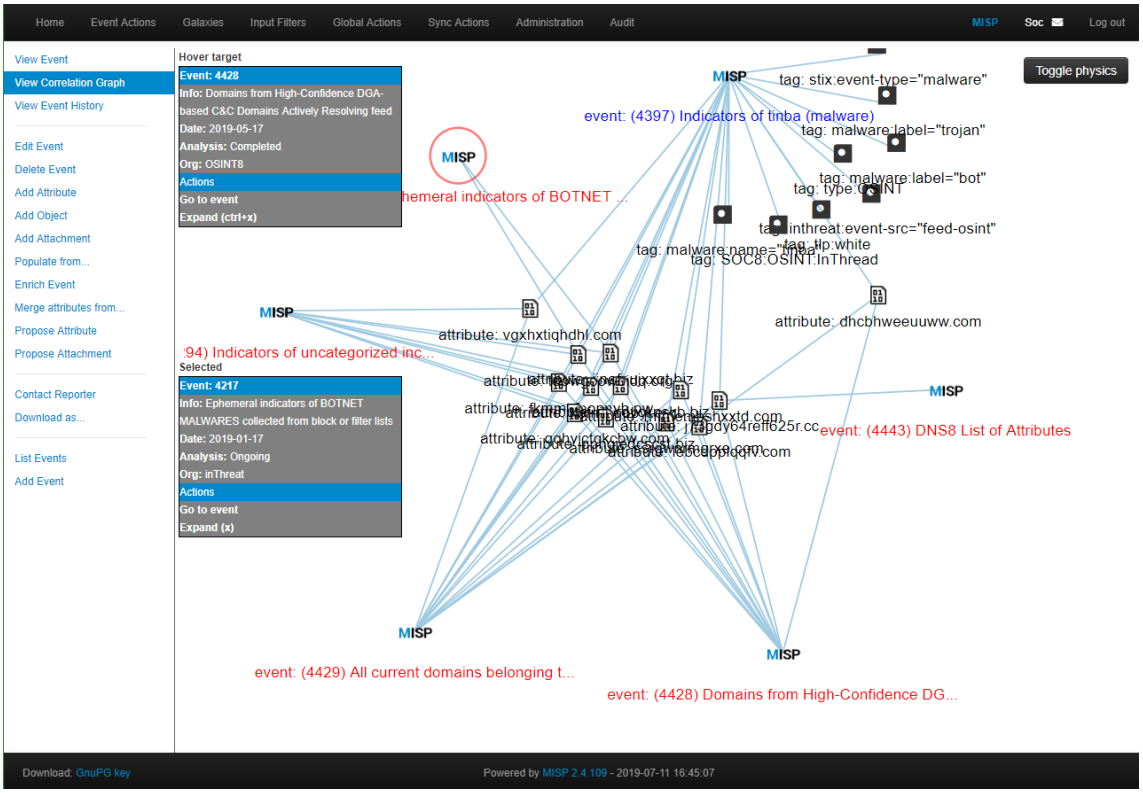

Figura 3.4: Correlação de eventos no MISP

3.2 Splunk

O Splunk[10] é a plataforma de SIEM que será utilizado para este projeto de dissertação. Esta plataforma é capaz de receber vários tipos de dados estruturados e não estruturados, normalizá-los, indexá-los e correlacionar os eventos em tempo-real. Com estes dados é

possível gerar gráficos, relatórios, alertas, *dashboards* e diferentes tipos de visualizações, como é possível verificar na figura 3.5.

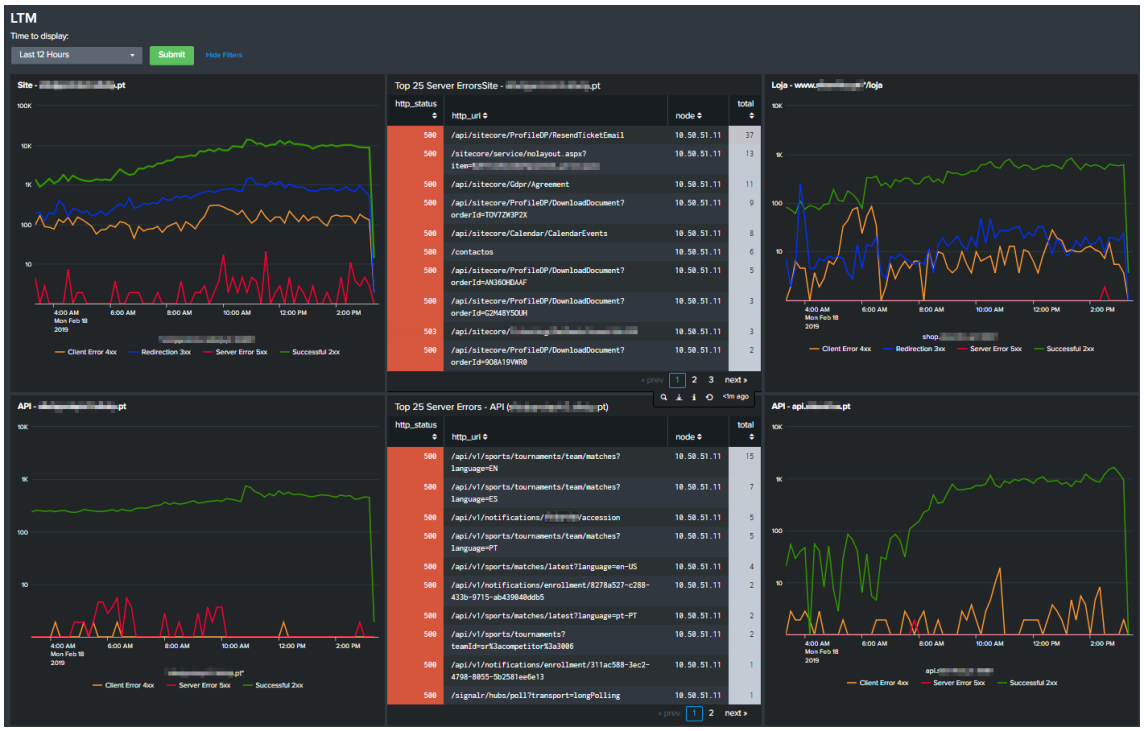

Figura 3.5: *Dashboards* de monitorização no Splunk

Esta plataforma é muito versátil, deixando de ser apenas uma plataforma para pesquisar dados num certo intervalo de tempo passando a ser uma aplicação que sustenta o negócio das empresas. As aplicações e integrações estão no centro da utilidade do Splunk sendo também bastante rápido e flexível a realizar as suas pesquisas.

O Splunk é uma ferramenta muito completa para a monitorização operacional, segurança e análise do comportamento do cliente. A monitorização operacional é possível devido à automação dos alertas e à investigação de incidentes pelos logs da infraestrutura. A nível de segurança devido a ser possível analisar todos os padrões de logs que não correspondem a padrões bons conhecidos. Com a junção de vários logs de vários sistemas, é possível assim compreender os padrões de cada cliente, por exemplo, saber que um certo utilizador costuma ligar-se à infraestrutura da organização via VPN por um certo país.

3.3 Cortex

O Cortex[6] é uma plataforma que tenta resolver problemas frequentemente comuns relacionados com *thread intel* bem como na resposta a incidentes pelas equipas de analistas. Esses problemas são nomeadamente a dificuldade em encontrar uma local central onde seja possível analisar um IOC de forma mais profunda. Atualmente existem muitas ferra-

mentas independentes onde é possível retirar várias informações acerca de um IOC, como é o caso do VirusTotal[12], AlienVault[27], Shodan[29] e AbuseIPDB[1].

Esta plataforma é uma ferramenta *open source* e agrupa diferentes analisador[2] de indicadores apenas neste local. Esses analisadores comunicam via API com os seus servidores, coletando assim várias informações à cerca de um IOC num único local, dando contexto e enriquecendo os indicadores. Isto possibilita ao analista de usar apenas esta ferramenta de uma forma rápida e eficiente para obter informações sobre um dado indicador.

Com este propósito é possível analisar observáveis, ou indicadores, tais como IPs, endereços de email, nomes de domínio, ficheiros, *hashes* e muito mais. Na figura 3.6 é possível ver um histórico de diferentes análises realizadas nesta plataforma.

O Cortex possuí também uma Rest API[7] que permite interagir com todas as funcionalidades desta plataforma, permitindo assim aumentar ainda mais o nível de automatização de qualquer processo de análise.

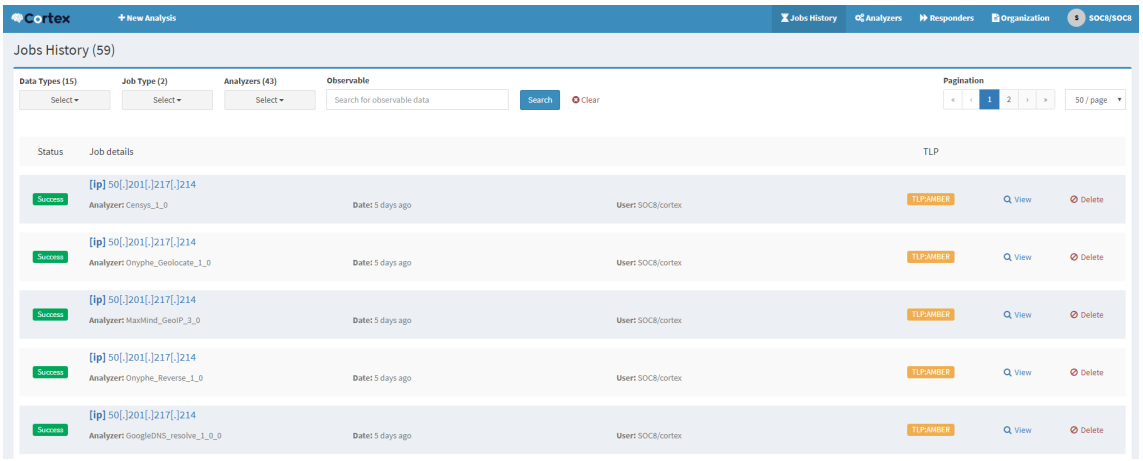

Figura 3.6: Hostórico de análises de indicadores no Cortex

3.4 DNS8 - LAYER8

[CONFIDENCIAL]

3.5 Conclusão

Neste capítulo foram detalhadas as plataformas essenciais que são utilizadas neste projeto de dissertação. O MISP e o DNS8 funcionam como as plataformas capazes de recolher indicadores de várias fontes, a plataforma Cortex é responsável pela análise de indicadores e o Splunk será a plataforma de SIEM já adotada pela LAYER8 anteriormente. Embora existam outras plataformas alternativas à solução, esta escolha foi a que melhor

se adequou para complementar as necessidades do serviço do SOC da LAYER8 com a elaboração deste projeto.

Em suma a tabela 3.5 representa o tipo de utilização nas plataformas a utilizar neste projeto de dissertação.

	SIEM	Recolha OSINT	Análise OSINT
MISP		X	
Splunk	X		
Cortex			X
DNS8		X	

Tabela 3.2: Representação das plataformas a utilizar e suas funcionalidades

Capítulo 4

Sistema Automático para Recolha de OSINT

[CONFIDENCIAL]

Capítulo 5

Módulo de Enriquecimento de Conhecimento de Indicadores

[CONFIDENCIAL]

Capítulo 6

Plataforma *Threat Intel*

[CONFIDENCIAL]

Capítulo 7

Integração com Plataforma *Threat Intel*

[CONFIDENCIAL]

Capítulo 8

Casos Práticos em Ambiente Real

[CONFIDENCIAL]

Capítulo 9

Conclusões

Neste capítulo final será descrito o resumo de tudo o que foi feito neste projeto de dissertação, seguindo de uma conclusão final com a avaliação do mesmo e de seguida propostas para trabalho futuro.

9.1 Conclusão

[CONFIDENCIAL]

9.2 Conclusão Final

A solução desta implementação e integração cumpriu com os objetivos propostos tendo sido possível recolher, armazenar e analisar indicadores, evitando que haja indicadores obsoletos, de forma a aumentar a confiança dos IOCs nos SIEMs, permitindo que o sistema tenha a capacidade de reconhecer e atuar contra novas ameaças em tempo útil.

O sistema desenvolvido encontra-se em produção não tendo resultado em nenhum produto comercial mas sim num complemento ao serviço do SOC da LAYER8. Este sistema acrescenta vantagens a nível de funcionalidades operacionais dos analistas do SOC e na automatização nas medidas de contenção proativas de incidentes de segurança dos seus clientes. Para além disto, este sistema contribuiu para a cooperação entre dois serviços da LAYER8, o SOC8 e o DNS8, através da integração destes dois serviços permitindo assim o crescimento dos serviços.

9.3 Trabalho Futuro

[CONFIDENCIAL]

Bibliografia

[1] AbuseIPDB. https://www.abuseipdb.com/, Acedido em: 2019-06-03.

[2] Analisadores cortex. https://github.com/TheHive-Project/Cortex-Analyzers, Acedido em: 07-12-2018.

[3] Biblioteca pymisp. https://github.com/CIRCL/PyMISP, Acedido em: 07-12-2018.

[4] Circl - *Computer IncidentResponse Center Luxembourg*. https://www.circl.lu/, Acedido em: 21-10-2018.

[5] Cncs - centro nacional de cibersegurança. https://www.cncs.gov.pt/, Acedido em: 21-02-2019.

[6] Cortex. https://github.com/TheHive-Project/Cortex, Acedido em: 27-11-2018.

[7] Cortex rest api. https://github.com/TheHive-Project/CortexDocs/blob/master/api/api-guide.md, Acedido em: 23-11-2018.

[8] Guia de partilha traffic light protocol (tlp). https://www.first.org/tlp/, Acedido em: 30-11-2018.

[9] Misp. https://www.misp-project.org/, Acedido em: 07-12-2018.

[10] Splunk. https://www.splunk.com/, Acedido em: 09-12-2018.

[11] Tallos intelligence. https://www.talosintelligence.com/, Acedido em: 05-02-2019.

[12] VirusTotal. https://www.virustotal.com/gui/home/upload, Acedido em: 2019-06-03.

[13] Taxonomia Comum para a Rede Nacional de CSIRTs. 2012.

[14] Amrit Williams. The Future of SIEM - The market will begin to diverge, 2007.

[15] Onur Catakoglu, Marco Balduzzi, and Davide Balzarotti. Automatic Extraction of Indicators of Compromise for Web Applications. pages 333–343, 2016.

[16] Da667. Instalação MISP - AutoMISP. `https://github.com/da667/AutoMISP`, Acedido em: 2018-09-30.

[17] David Nathans. *Designing and Building Security Operations Center*, volume 1. 2014.

[18] Kai Oliver Detken, Thomas Rix, Carsten Kleiner, Bastian Hellmann, and Leonard Renners. SIEM approach for a higher level of IT security in enterprise networks. *Proceedings of the 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2015*, 1(September):322–327, 2015.

[19] Roger Dingledine, Nick Mathewson, and Paul Syverson. Dingledine et al. - Tor - the second-generation onion router - 2004.

[20] Diogo Carou. Revisiting RFC2350 20 Years Later: A Hands-On Approach to Security Monitoring and Incident Response. Master's thesis, Universidade de Lisboa, 2018.

[21] Fulano, Cicrano, and Beltrano. A paper on something. In *The 7th Conference on Things and Stuff (CTS 2009)*, Lisbon, Portugal, May 2009. Accepted for publication.

[22] LAYER8. LAYER8. `https://www.layer8.pt/`, Acedido em: 2019-06-03.

[23] LAYER8. DNS8. 2015.

[24] LAYER8. Type of Incidents and Classifications. Technical report, 2015.

[25] LAYER8. *SOC8 Incident Taxonomy*. LAYER8, Lisboa, 2017.

[26] MISP-Taxonomia. Taxonomia MISP. `https://github.com/MISP/misp-taxonomies`, Acedido em: 04-06-2019.

[27] OTXAlienVault. OTX AlienVault. `https://otx.alienvault.com/`, Acedido em: 2019-06-03.

[28] Iztok Podbregar. International Journal of Intelligence and OSINT : A " Grey Zone "? (May), 2014.

[29] Shodan. Shodan. `https://www.shodan.io/`, Acedido em: 2019-06-03.

[30] Stanislav Spacek, Martin Lastovicka, Martin Horak, and Tomas Plesnik. Current Issues of Malicious Domains Blocking. pages 551–556.

[31] P. Vixie and V. Schryver. DNS Response Policy Zones (RPZ). `https://tools.ietf.org/html/draft-vixie-dnsop-dns-rpz-00`, Acedido em: 09/06/2019.

[32] Carson Zimmerman. *Cybersecurity Operations Center*. 2014.